Gebze Technical University

CSE455/CSE552 – Machine Learning (Spring 2023)

KNN, SVM and DT in a Classification & Regression

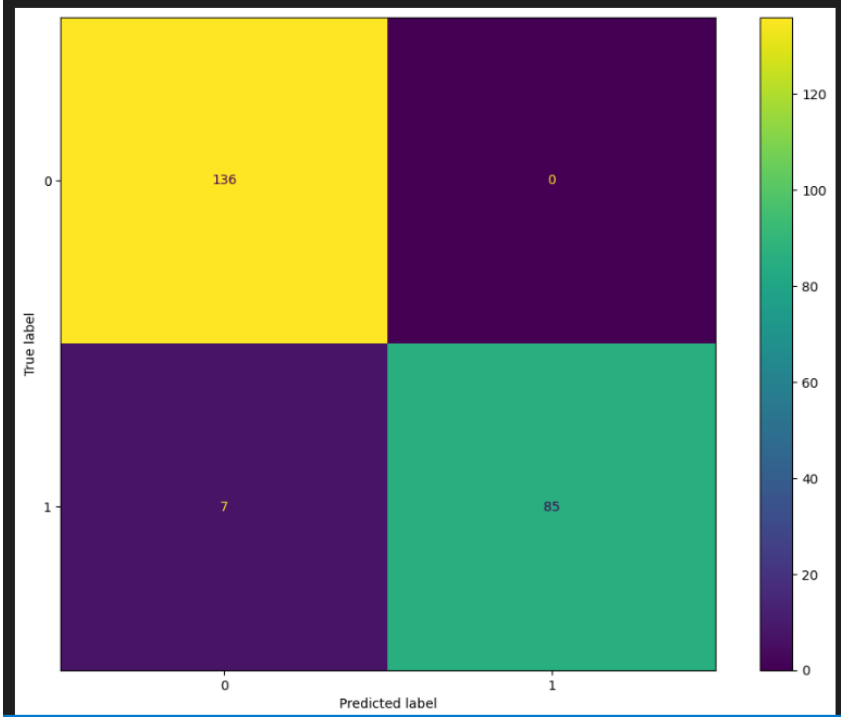
Homework#1 Report

Muhammed Bedir ULUCAY

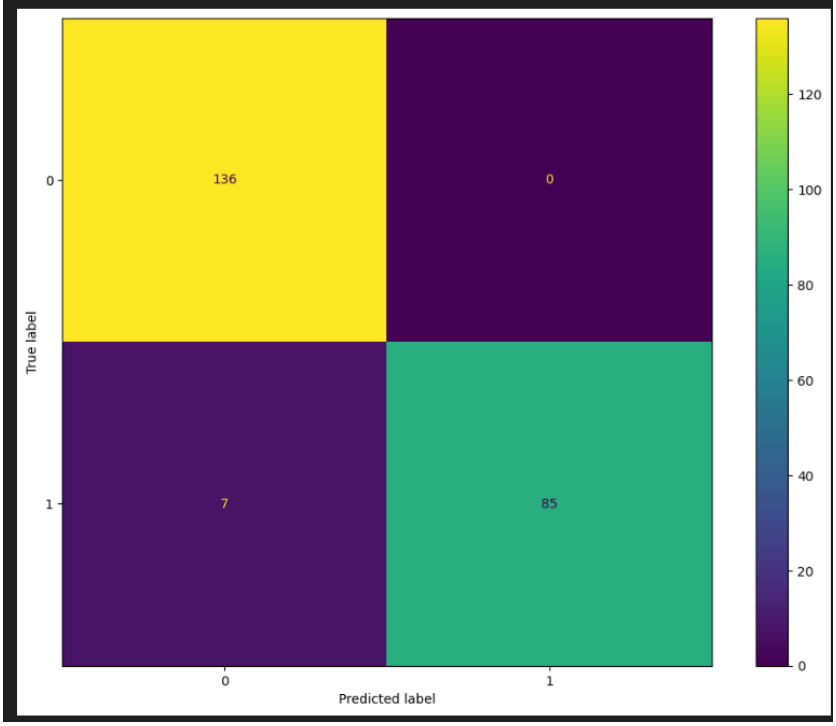
1901042697

PART 1:

```
manhattan_distance
Accuracy : % 96.9298245614035
              precision    recall  f1-score   support

         0.0       0.95      1.00      0.97       136
         1.0       1.00      0.92      0.96        92

    accuracy                           0.97       228
   macro avg       0.98      0.96      0.97       228
weighted avg       0.97      0.97      0.97       228

[[136   0]
 [  7  85]]
```



```
euclidean_distance
Accuracy : % 96.9298245614035
              precision    recall  f1-score   support

         0.0       0.95      1.00      0.97       136
         1.0       1.00      0.92      0.96        92

    accuracy                           0.97       228
   macro avg       0.98      0.96      0.97       228
weighted avg       0.97      0.97      0.97       228

[[136   0]
 [  7  85]]
```

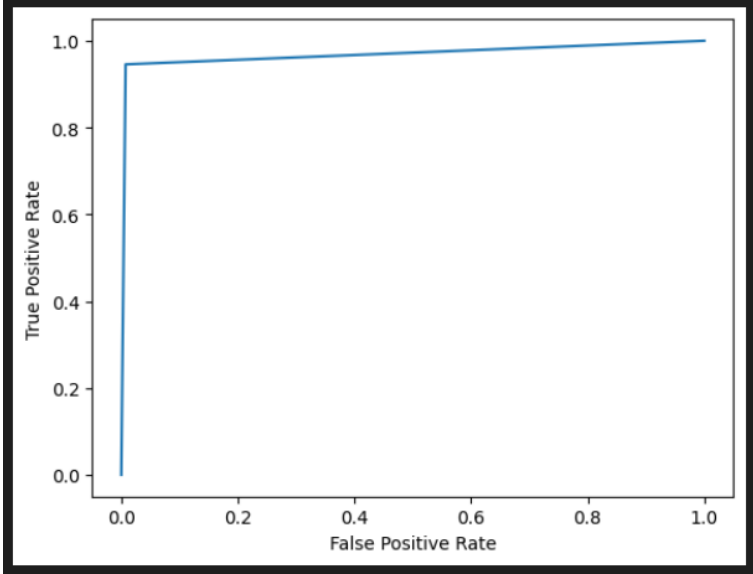



```
CPU times: total: 656 ms
Wall time: 649 ms
```

PART 2:

```
Regression score is: 0.9939 for k_value: 4
Original Model

n_neighbors: 9

R2: 0.99
Adj_R2: 0.99
Mean Absolute Error: 9.61
Mean Squared Error: 185.06
Root Mean Squared Error: 13.60
CPU times: total: 2.55 s
Wall time: 2.54 s
```

PART 3:




```
[2. 1. 0.]
Accuracy : % 97.36842105263158
              precision    recall  f1-score   support

         0.0       0.96      0.99      0.98       136
         1.0       0.99      0.95      0.97        92

    accuracy                           0.97       228
   macro avg       0.98      0.97      0.97       228
weighted avg       0.97      0.97      0.97       228

CPU times: total: 62.5 ms
Wall time: 70 ms
```

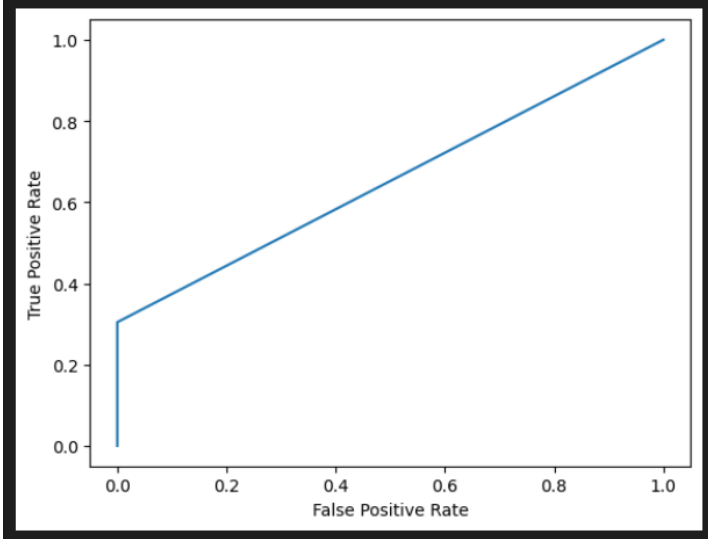
PART 4:

```
Regression score is: 0.9999
CPU times: total: 58.9 s
Wall time: 59.5 s
```

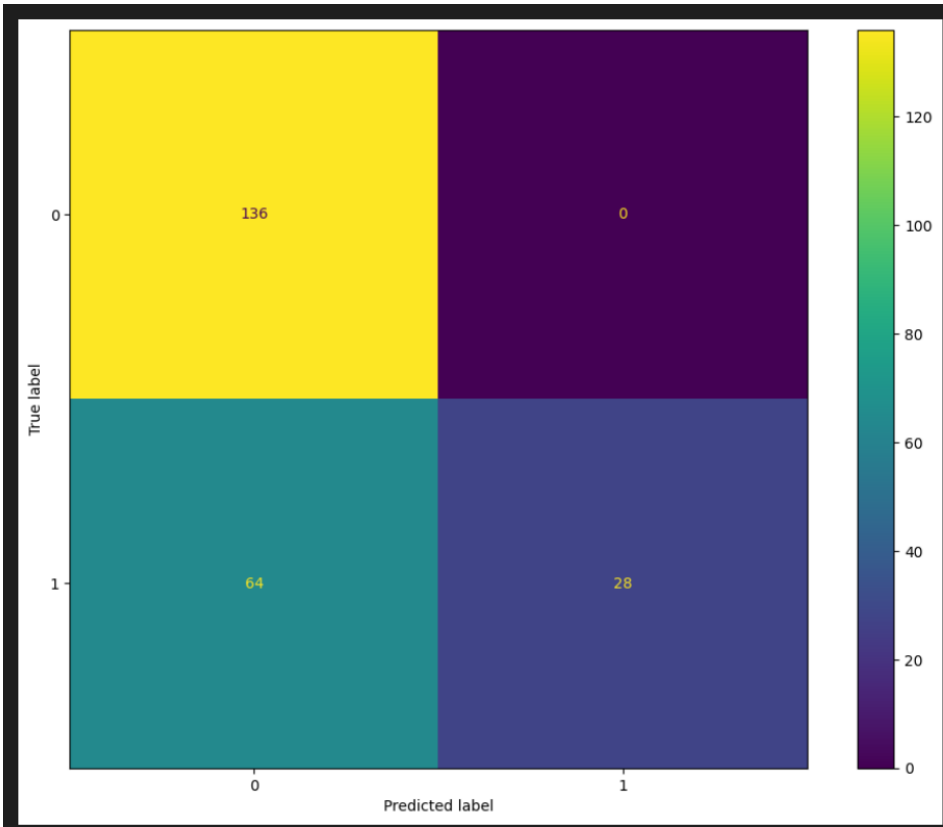
```
SVR Regression Score (mean) for train set: 1.0000 , Standard deviation for train set: 0.0000
SVR Regression Score (mean) for test set: 1.0000 , Standard deviation for test set: 0.0000
CPU times: user 11min 58s, sys: 931 ms, total: 11min 59s
Wall time: 12min 6s
```

PART 5:

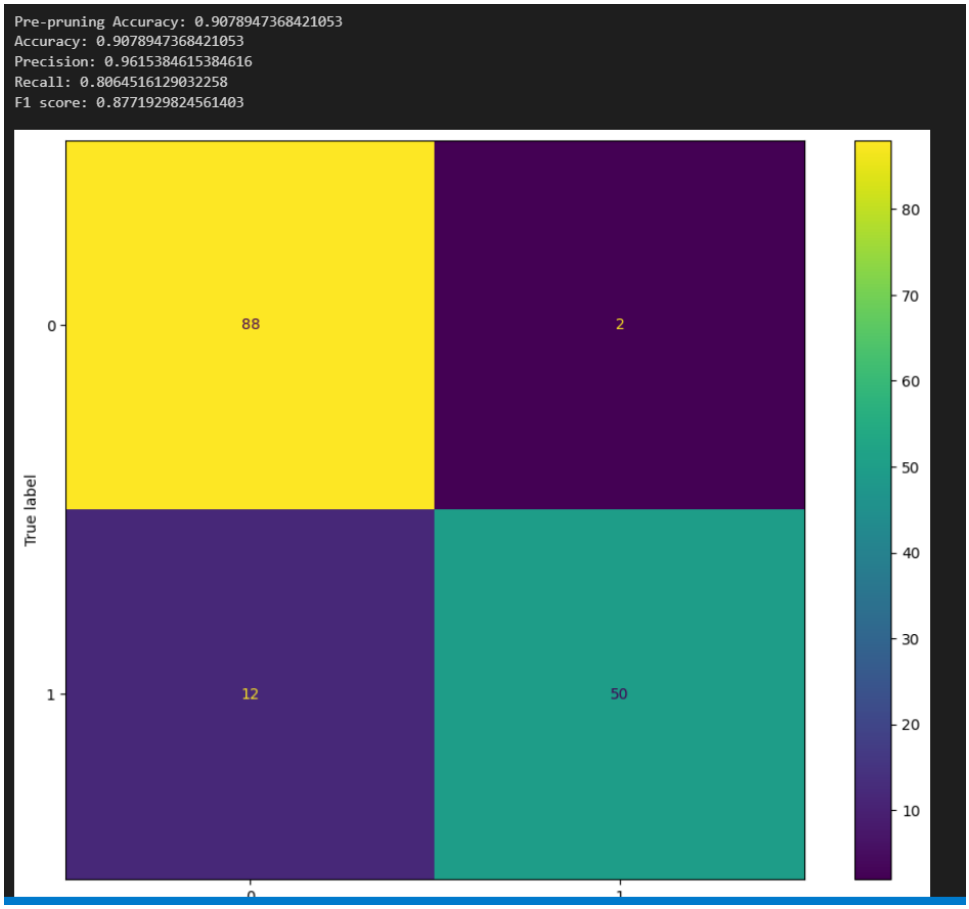
```
Accuracy: 0.7192982456140351
Original Model
```



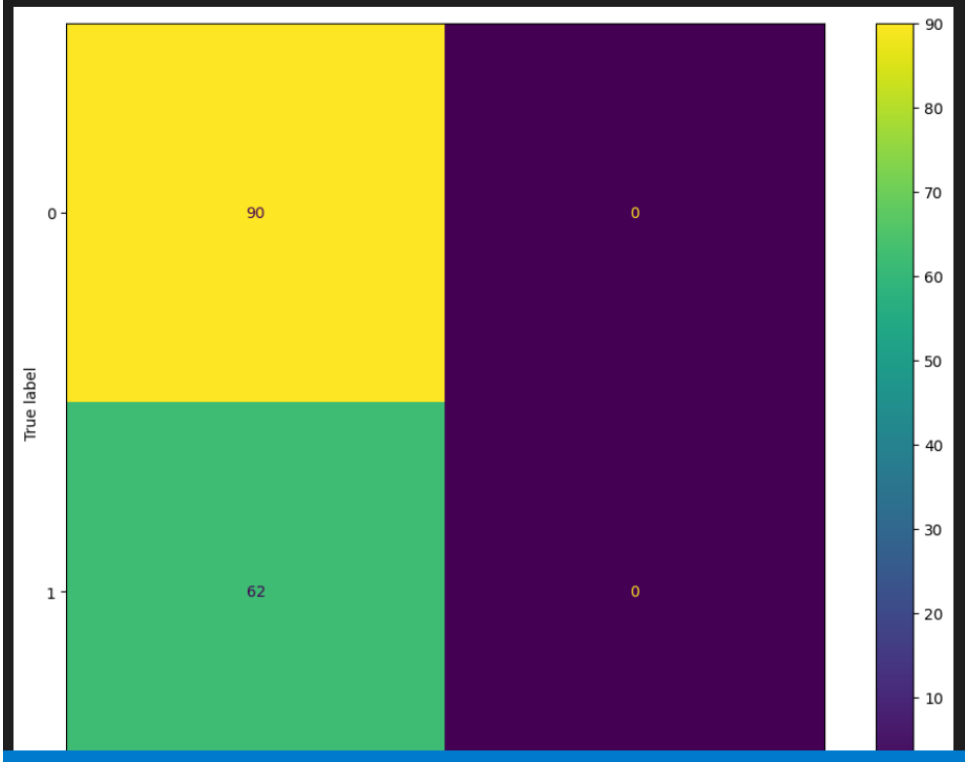

```
[2. 1. 0.]
              precision    recall  f1-score   support

         0.0       0.68      1.00      0.81       136
         1.0       1.00      0.30      0.47        92

    accuracy                           0.72       228
   macro avg       0.84      0.65      0.64       228
weighted avg       0.81      0.72      0.67       228
```



PART 6:


Pre-pruning Accuracy: 0.9078947368421053
Accuracy: 0.9078947368421053
Precision: 0.9615384615384616
Recall: 0.8064516129032258
F1 score: 0.8771929824561403
88
2
12
50
True label
0
1
0
1
80
70
60
50
40
30
20
10

```
Post-pruning Accuracy: 0.5921052631578947
Accuracy: 0.5921052631578947
Precision: 0.0
Recall: 0.0
F1 score: 0.0
C:\Users\kral_\AppData\Local\Packages\PythonSoftwareFoundation.Python.3.10_qbz5n2kfra8p0\LocalCache\local-package
  _warn_prf(average, modifier, msg_start, len(result))
```




```
([('Score', '<=', 2.3000000715255737)], 0)
([('Score', '>', 2.3000000715255737), ('Risk_D', '<=', 0.4220000058412552), ('CONTROL_RISK', '<=', 0.6000000089406967), ('Risk_A', '<=', 3.615000009536743), ('Risk_B', '<=', 3
([('Score', '>', 2.3000000715255737), ('Risk_D', '<=', 0.4220000058412552), ('CONTROL_RISK', '<=', 0.6000000089406967), ('Risk_A', '<=', 3.615000009536743), ('Risk_B', '<=', 3
([('Score', '>', 2.3000000715255737), ('Risk_D', '<=', 0.4220000058412552), ('CONTROL_RISK', '<=', 0.6000000089406967), ('Risk_A', '<=', 3.615000009536743), ('Risk_B', '<=', 3
([('Score', '>', 2.3000000715255737), ('Risk_D', '<=', 0.4220000058412552), ('CONTROL_RISK', '<=', 0.6000000089406967), ('Risk_A', '<=', 3.615000009536743), ('Risk_B', '<=', 3
([('Score', '>', 2.3000000715255737), ('Risk_D', '<=', 0.4220000058412552), ('CONTROL_RISK', '<=', 0.6000000089406967), ('Risk_A', '<=', 3.615000009536743), ('Risk_B', '>', 3.
([('Score', '>', 2.3000000715255737), ('Risk_D', '<=', 0.4220000058412552), ('CONTROL_RISK', '<=', 0.6000000089406967), ('Risk_A', '>', 3.615000009536743)], 1)
([('Score', '>', 2.3000000715255737), ('Risk_D', '<=', 0.4220000058412552), ('CONTROL_RISK', '>', 0.6000000089406967), ('CONTROL_RISK', '<=', 1.0000000298023224), ('Risk_B', '
([('Score', '>', 2.3000000715255737), ('Risk_D', '<=', 0.4220000058412552), ('CONTROL_RISK', '>', 0.6000000089406967), ('CONTROL_RISK', '<=', 1.0000000298023224), ('Risk_B', '
([('Score', '>', 2.3000000715255737), ('Risk_D', '<=', 0.4220000058412552), ('CONTROL_RISK', '>', 0.6000000089406967), ('CONTROL_RISK', '>', 1.0000000298023224)], 1)
([('Score', '>', 2.3000000715255737), ('Risk_D', '>', 0.4220000058412552), ('Risk_D', '<=', 0.7949999868869781), ('Risk_D', '<=', 0.7609999775886536), ('Risk_B', '<=', 2.35499
([('Score', '>', 2.3000000715255737), ('Risk_D', '>', 0.4220000058412552), ('Risk_D', '<=', 0.7949999868869781), ('Risk_D', '<=', 0.7609999775886536), ('Risk_B', '<=', 2.35499
([('Score', '>', 2.3000000715255737), ('Risk_D', '>', 0.4220000058412552), ('Risk_D', '<=', 0.7949999868869781), ('Risk_D', '<=', 0.7609999775886536), ('Risk_B', '>', 2.354999
([('Score', '>', 2.3000000715255737), ('Risk_D', '>', 0.4220000058412552), ('Risk_D', '<=', 0.7949999868869781), ('Risk_D', '>', 0.7609999775886536)], 0)
([('Score', '>', 2.3000000715255737), ('Risk_D', '>', 0.4220000058412552), ('Risk_D', '>', 0.7949999868869781), ('Score', '<=', 2.5), ('MONEY_Marks', '<=', 5.0), ('Risk_D', '<
([('Score', '>', 2.3000000715255737), ('Risk_D', '>', 0.4220000058412552), ('Risk_D', '>', 0.7949999868869781), ('Score', '<=', 2.5), ('MONEY_Marks', '<=', 5.0), ('Risk_D', '>
([('Score', '>', 2.3000000715255737), ('Risk_D', '>', 0.4220000058412552), ('Risk_D', '>', 0.7949999868869781), ('Score', '<=', 2.5), ('MONEY_Marks', '>', 5.0)], 1)
([('Score', '>', 2.3000000715255737), ('Risk_D', '>', 0.4220000058412552), ('Risk_D', '>', 0.7949999868869781), ('Score', '>', 2.5)], 1)
Accuracy: 0.9802631578947368
```

PART 7:

```
Accuracy: 0.9735327963176065
Precision: 0.9731396240889911
Recall: 0.9735327963176065
F1 score: 0.9720779220779221
```